An Augmented Lagrange Method for Ill-Posed Elliptic State Constrained Optimal Control Problems with Sparse Controls

Veronika Karl, Frank Pörner

An augmented Lagrange method for ill-posed elliptic state constrained optimal control problems with sparse controls[*]

Veronika Karl[†]   Frank Pörner[‡]

July 26, 2017

Abstract

We provide a modified augmented Lagrange method for solving ill-posed state-constrained elliptic optimal control problems with sparse controls. We consider a linear quadratic optimal control problem without any additional L^2 regularization terms. The sparsity is guaranteed by an additional L^1 term. Here, the modification of the classical augmented Lagrange method guarantees us uniform boundedness of the multiplier that corresponds to the state constraints as well as strong convergence of the control and its corresponding state. Moreover convergence results proving the weak convergence of the adjoint state and weak*-convergence of the multiplier are provided. Finally, we demonstrate our method in several numerical examples.

Keywords: ill-posed optimal control, state constraints, augmented Lagrange method.

AMS subject classification: 49M20, 65K10, 90C30.

1   Introduction

In this paper we consider a convex optimal control problem of the following form

$$\min\ J(y,u) := \frac{1}{2}||y - y_d||_{L^2(\Omega)}^2 + \beta\|u\|_{L^1(\Omega)} \qquad (P)$$

subject to

$$
\begin{aligned}
Ay &= u & &\text{in } \Omega, \\
y &= 0 & &\text{on } \partial\Omega, \\
y &\leq \psi & &\text{in } \Omega, \\
u_a \leq u &\leq u_b & &\text{in } \Omega.
\end{aligned}
$$

We set $j(u) := \|u\|_{L^1(\Omega)}$ for abbreviation. Here A is a linear elliptic operator and $\beta \geq 0$. The main difficulties in this problem are the pointwise state constraints $y(x) \leq \psi(x)$ and the convex but non-differentiable term $\|u\|_{L^1(\Omega)}$. Note that there is no additional L^2 regularization term present in (P) which makes the problem ill-posed and numerically challenging. The motivation for the L^1-term in the cost functional is the following. The optimal solution $\bar{u}$ of (P) is sparse, i.e., the control is zero on large parts of the domain if β is large enough. This can be used in the

optimal placement of controllers, especially in situations where it is not desirable to control the system from the whole domain Ω, see [30]. Such sparsity promoting optimal control problems without state constraints have been studied in, e.g. [32–34] for optimal control of linear partial differential equations and in [6, 8] for the optimal control of semilinear equations. For sufficient second-order conditions for the state constrained sparsity promoting optimal control problem with a semilinear partial differential equation we refer to [10].

In order to deal with the state constraints we apply an augmented Lagrange method established by the first author in [22]. There the optimal control problem

$$\text{Minimize } \frac{1}{2}\|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2}\|u\|_{L^2(\Omega)}^2$$

with $\alpha > 0$ subject to an elliptic linear partial differential equation, state constraints and bilateral control constraints had been considered. Under suitable regularity assumptions the existence of Lagrange multipliers can be proven. However in many cases the multiplier $\bar{\mu}$ has a very low regularity, e.g. $\bar{\mu} \in C(\Omega)^* = \mathcal{M}(\bar{\Omega})$, where $\mathcal{M}(\bar{\Omega})$ denotes the space of regular Borel measures on $\bar{\Omega}$. This makes the numerical solution of (P) very challenging. Augmented Lagrange methods are well-known in optimization. In [1, 2] the state equation is augmented, and in [19] they deal with finitely many state constraints.

Our aim is to modify and extend the method presented in [22] to obtain a numerical scheme to solve (P). The main idea is the following. We add a Tikhonov regularization term $\frac{\alpha}{2}\|u\|_{L^2(\Omega)}^2$ to (P) and apply the augmented Lagrange method. During the algorithm we decrease the regularization parameter $\alpha \to 0$. Since this is a classical Tikhonov regularization approach, we aim to achieve strong convergence against the solution of (P).

This approach is similar to the proximal method of multipliers [28] for finite dimensions. There in each iterations a regularization term $\|u_k - u\|_{L^2(\Omega)}^2$ is added. However this is not a suitable choice here, since our method reduces to the proximal point method [13, 15, 21, 25, 29] if the state constraints are not active. Due to the example given by Güler [14] the iterates generated by the proximal point method are in general only weakly converging.

The paper is structured as follows. First we recall in section 2 some preliminary results, then we analyze the Tikhonov regularization of (P) in section 3. The augmented Lagrange method will be introduced in section 4. Similar to [22] we only update the multiplier if a certain measure of feasibility and violation of complementarity shows sufficient decrease. In section 5 we establish convergence of our method, which is the main result of this paper. The implementation our the algorithm is described in section 6 and numerical results are be presented in 7.

Notation. Throughout the article we will use the following notation. The inner product in $L^2(\Omega)$ is denoted by $(\cdot, \cdot)$. Duality pairings will be denoted by $\langle \cdot, \cdot \rangle$. The dual of $C(\bar{\Omega})$ is denoted by $\mathcal{M}(\bar{\Omega})$, which is the space of regular Borel measures on $\bar{\Omega}$. Furthermore c is a generic constant which may change from line to line, but is independent from the important variables, e.g. k.

2 Preliminary results

2.1 Problem setting

Let $\Omega \subset \mathbb{R}^N$, $N = \{1, 2, 3\}$ be a bounded domain with $C^{0,1}$-boundary Γ. Let Y denote the space $Y := H_0^1(\Omega) \cap C(\bar{\Omega})$ and $U := L^2(\Omega)$. We want to solve the following state constrained optimal control problem: Minimize

$$J(y, u) := \frac{1}{2}\|y - y_d\|_{L^2(\Omega)}^2 + \beta\|u\|_{L^1(\Omega)}$$

over all $(y, u) \in Y \times U$ subject to the elliptic equation

$$(Ay)(x) = u(x) \quad \text{in } \Omega,$$
$$y(x) = 0 \qquad \text{on } \Gamma,$$

and subject to the pointwise state and control constraints

$$y(x) \leq \psi(x) \qquad \text{in } \Omega,$$
$$u_a(x) \leq u(x) \leq u_b(x) \quad \text{a.e. in } \Omega.$$

In the sequel, we will work with the following set of standing assumptions.

Assumption 1. *1. The given data satisfy $y_d \in L^2(\Omega)$, $u_a, u_b \in L^\infty(\Omega)$ with $u_a \leq 0 \leq u_b$ and $\psi \in C(\bar\Omega)$.*

  2. The differential operator A is given by

$$(Ay)(x) := -\sum_{i,j=1}^{N} \partial_{x_j}(a_{ij}(x)\partial_{x_i}y(x))$$

  with $a_{i,j} \in C^{0,1}(\bar\Omega)$. The operator A is assumed to be strongly elliptic, i.e., there is $\delta > 0$ such that

$$\sum_{i,j=1}^{N} a_{ij}(x)\xi_i\xi_j \geq \delta|\xi|^2 \quad \forall \xi \in \mathbb{R}^N, \ \text{a.e. on } \Omega.$$

The following theorem is taken from [7, Theorem 2.1].

Theorem 2.1. *For every $u \in L^2(\Omega)$ there exists a unique weak solution $y \in H_0^1(\Omega) \cap C(\bar\Omega)$ of the state equation and it holds*

$$\|y\|_{H_0^1(\Omega)} + \|y\|_{C(\bar\Omega)} \leq c \|u\|_{L^2(\Omega)}.$$

with a constant $c > 0$ independent of u.

With this assumption one can prove the following properties of the control-to-state mapping S.

Theorem 2.2. *The control-to-state operator $S : L^2(\Omega) \to H_0^1(\Omega) \cap C(\bar\Omega), u \mapsto y$ is a linear, continuous, and compact operator.*

Proof. The linearity follows directly by the definition of S and for the compactness we refer [7, Theorem 2.1]. □

In the following, we will use the feasible sets with respect to the state and control constraints denoted by

$$U_{\text{ad}} = \{u \in L^\infty(\Omega) \mid u_a(x) \leq u(x) \leq u_b(x) \text{ a.e. in } \Omega\},$$
$$Y_{\text{ad}} = \{y \in C(\bar\Omega) \quad \mid y(x) \leq \psi(x) \ \forall x \in \Omega\}.$$

The feasible set of the optimal control problem is denoted by

$$F_{\text{ad}} = \{(y, u) \in Y \times L^2(\Omega) \mid (y, u) \in Y_{\text{ad}} \times U_{\text{ad}}, \ y = Su\}.$$

The assumption $u_a \leq 0 \leq u_b$ is not a restriction. Assume that $u_a > 0$ on a subset $\Omega_1 \subseteq \Omega$. Then we can decompose the L^1-norm for $u \in U_{\text{ad}}$ as $\|u\|_{L^1(\Omega)} = \|u\|_{L^1(\Omega \backslash \Omega_1)} + \int_{\Omega_1} u$. Hence, on Ω_1 the L^1-norm is a linear functional and its treatment does not impose any further difficulties.

Theorem 2.3. *Assume that the feasible set F_{ad} is non-empty. Then, there exists a unique solution $\bar u$ with associated state $\bar y$ of (P).*

Proof. The existence of solutions follows by standard arguments. Due to the assumptions the operator S is linear, continuous, and injective. Hence the problem (P) is convex leading to a unique optimal state $\bar y$. By using the injectivity of S we now obtain uniqueness of the optimal control. □

2.2 Subdifferential of j

In this section we want to recall some basic properties of the subdifferential of the function $j(u) = \|u\|_{L^1(\Omega)}$. Since j is convex and Lipschitz, the generalized gradient (see [11]) and the subdifferential in the sense of convex analysis coincide. The subdifferential is defined by

$$\partial j(u) := \left\{ \lambda \in L^\infty(\Omega) : \int_\Omega \lambda(v-u)\,\mathrm{d}x \leq \|v\|_{L^1(\Omega)} - \|u\|_{L^1(\Omega)}, \quad \forall v \in L^1(\Omega) \right\}.$$

Since j is a convex function with $\mathrm{dom}(j) = L^1(\Omega)$ the subdifferential is always nonempty. It is easy to compute that $\lambda \in \partial j(u)$ if and only if

$$\lambda \begin{cases} = +1 & \text{if } u(x) > 0 \\ = -1 & \text{if } u(x) < 0 \\ \in [-1, +1] & \text{if } u(x) = 0 \end{cases}.$$

For more information we refer to the book of Bonnans and Shapiro [5, Section 2.4.3]. We will need the subdifferential to establish derivatives of the objective functional $\frac{1}{2}\|y - y_d\|^2 + \beta\|u\|_{L^1(\Omega)}$ and to obtain optimality conditions.

2.3 Optimality conditions

The existence of Lagrange multipliers cannot be guaranteed without any further regularity assumptions. Throughout this paper will assume that the following Slater condition is satisfied.

Assumption 2. *We assume that there exists $\hat{u} \in U_{\mathrm{ad}}$ and $\sigma > 0$ such that for $\hat{y} = S\hat{u}$ it holds*

$$\hat{y}(x) \leq \psi(x) - \sigma \quad \forall x \in \Omega.$$

The choice of Assumption 2 as regularity condition is motivated as follows. The given inequality in the Slater condition coincides with $\psi - \hat{y}$ lying in the interior of the nonnegative cone of Y. The nonnegative cone of $Y = H_0^1(\Omega) \cap C(\bar{\Omega})$ equipped with its natural norm $\|\cdot\|_Y := \|\cdot\|_{H_0^1(\Omega)} + \|\cdot\|_{C(\Omega)}$ has nonempty interior - in contrast to $L^p(\Omega)$, $p \in [1, \infty)$ equipped with the L^p-norm. This implies a possible existence of a Slater point $\hat{u}$ that satisfies Assumption 2. Moreover, since S is linear, Assumption 2 is equivalent to the linearized Slater condition, which on the other hand implies the more general Zowe-Kurcyusz regularity condition (see [31, p.332]). However, since the set of feasible controls may have no interior points (for an example see [31]), the Zowe-Kurcyusz regularity condition does not imply the linearized Slater condition. Furthermore, one already has to know the solution of the optimal control problem (P) to check whether the Zowe-Kurcyusz condition is satisfied. This is not the case for the proposed Slater condition.

Theorem 2.4. *Let $(\bar{u}, \bar{y})$ be a solution of the problem (P). Furthermore, let Assumption 2 be fulfilled. Then, there exists an adjoint state $\bar{p} \in W_0^{1,s}(\Omega)$, $s \in [1, N/(N-1))$, a Lagrange multiplier $\bar{\mu} \in \mathcal{M}(\bar{\Omega})$ and a subdifferential $\bar{\lambda} \in \partial j(\bar{u})$ such that the following optimality system*

$$\begin{cases} A\bar{y} = \bar{u} & \text{in } \Omega, \\ \bar{y} = 0 & \text{on } \Gamma, \end{cases} \tag{1a}$$

$$\begin{cases} A^*\bar{p} = \bar{y} - y_d + \bar{\mu} & \text{in } \Omega, \\ \bar{p} = 0 & \text{on } \Gamma, \end{cases} \tag{1b}$$

$$(\bar{p} + \beta\bar{\lambda}, u - \bar{u}) \geq 0 \quad \forall u \in U_{ad}, \tag{1c}$$

$$\langle \bar{\mu}, \bar{y} - \psi \rangle_{\mathcal{M}(\bar{\Omega}), C(\bar{\Omega})} = 0, \quad \bar{\mu} \geq 0, \tag{1d}$$

is fulfilled. Here, the inequality $\bar{\mu} \geq 0$ means $\langle \bar{\mu}, \varphi \rangle_{\mathcal{M}(\bar{\Omega}), C(\bar{\Omega})} \geq 0$ for all $\varphi \in C(\bar{\Omega})$ with $\varphi \geq 0$.

Proof. The proof can be found in [10, Theorem 2.5]. □

In the definition for the optimal adjoint state $\bar{p}$ we have to solve an elliptic equation with a measure on the right hand side. This problem is well-posed in the following sense.

Theorem 2.5. *Let $\bar{\mu} \in \mathcal{M}(\bar{\Omega})$ be a regular Borel measure. Then the adjoint state equation*

$$A^*\bar{p} = \bar{y} - y_d + \bar{\mu} \quad in \ \Omega,$$
$$\bar{p} = 0 \qquad\qquad on \ \Gamma$$

has a unique very weak solution $\bar{p} \in W_0^{1,s}(\Omega)$, $s \in [1, N/(N-1))$, and it holds

$$\|\bar{p}\|_{W_0^{1,s}(\Omega)} \leq c \left(\|\bar{y}\|_{L^2(\Omega)} + \|y_d\|_{L^2(\Omega)} + \|\bar{\mu}\|_{\mathcal{M}(\bar{\Omega})} \right). \tag{2}$$

Proof. This result is due to [9, Theorem 4]. □

The next theorem shows the relation between the adjoint state and the control. One can see, that if β is large, the control will be zero on large parts of Ω. Hence $\bar{u}$ is sparse.

Lemma 2.6. *Let $\bar{u}, \bar{p}, \bar{\lambda}, \bar{\mu}$ satisfy the optimality system. (1a)-(1d). Then the following relations hold for $\theta > 0$:*

$$\bar{u}(x) \begin{cases} = u_a(x) & if \quad \bar{p}(x) > \beta \\ = u_b(x) & if \quad \bar{p}(x) < -\beta \\ = 0 & if \quad |p(x)| < \beta \\ \in [u_a(x), u_b(x)] & if \quad |p(x)| = \beta \end{cases},$$

$$\bar{\lambda}(x) = P_{[-1,+1]}\left(-\frac{1}{\beta}\bar{p}(x) \right),$$

$$\bar{u}(x) = P_{[u_a(x),u_b(x)]}\big(\bar{u}(x) - \theta(\bar{p}(x) + \beta\bar{\lambda}(x)) \big).$$

From the second formula it follows that $\bar{\lambda}$ is unique if the multiplier $\bar{\mu}$ and adjoint state $\bar{p}$ are unique.

Proof. The proof only uses the optimality (1c) and can be found in [6, Theorem 3.1]. □

3 The regularized problem

Solving the problem (P) directly is challenging for mainly two reasons. First, since the multiplier corresponding to the state constraints appears in form of a measure, it is not clear how to deal with the state constraints. For the control constraints many powerful methods are available. Here, we only want to mention the semi-smooth Newton solvers [16, 17] and the Active-Set methods [3]. However it is not clear how to implement the state constraints into a direct solver. In [4, 20] Active-Set methods has been used to solve problems where the state constraints have been treated by Moreau-Yosida regularization. In [20] also relations between semi-smooth Newton methods and Active-Set methods have been established that can be used to prove fast local convergence. In this work we want to adapt the approach of a modified augmented Lagrange method that has been proposed by the first author in [22] to overcome the lack of the multiplier's regularity.

The second challenge is the ill-posedness of the original problem (P). There small perturbations of the given data y_d may lead to large errors in the associated optimal controls. To deal with this issue we will use the well-known Tikhonov regularization technique with some positive regularization parameter $\alpha > 0$. The regularized problem is given by

$$\text{Minimize } J_\alpha(y, u) := \frac{1}{2}\|y - y_d\|_{L^2(\Omega)}^2 + \beta\|u\|_{L^1(\Omega)} + \frac{\alpha}{2}\|u\|_{L^2(\Omega)}^2$$

$$\begin{aligned} s.t. \quad & Ay = u \quad in \ \Omega, \\ & y = 0 \quad on \ \partial\Omega, \\ & y \leq \psi, \\ & u \in U_{\mathrm{ad}}. \end{aligned} \tag{P^α}$$

It is clear that (P^α) omits a unique solution u^α with associated state y^α. One can expect that u^α converges to the solution of (P) as $\alpha \to 0$. Similar results can be found in the literature, e.g. [32].

Lemma 3.1. *Let u^α be the unique solution of (P^α) with $\alpha > 0$. Furthermore let $\bar{u}$ be the unique solution of (P). Then we have $u^\alpha \to \bar{u}$ in $L^2(\Omega)$ as $\alpha \to 0$.*

Proof. We first show that $\|u^\alpha\|_{L^2(\Omega)} \leq \|\bar{u}\|_{L^2(\Omega)}$ for all $\alpha > 0$. Let J_0 denote the cost functional J_α for $\alpha := 0$. We start with

$$J_0(u^\alpha) + \frac{\alpha}{2}\|u^\alpha\|^2_{L^2(\Omega)} = J_\alpha(u^\alpha) \leq J_\alpha(\bar{u}) = J_0(\bar{u}) + \frac{\alpha}{2}\|\bar{u}\|^2_{L^2(\Omega)} \leq J_0(u^\alpha) + \frac{\alpha}{2}\|\bar{u}\|^2_{L^2(\Omega)},$$

where we exploited the optimality of u^α for (P^α) and the optimality of $\bar{u}$ for (P). This yields $\|u^\alpha\|_{L^2(\Omega)} \leq \|\bar{u}\|_{L^2(\Omega)}$. Now we use that the set U_{ad} is weakly compact and extract a subsequence $u^{\alpha_i} \rightharpoonup u^* \in U_{\mathrm{ad}}$. Since the operator S is compact, see Theorem 2.2, we obtain strong convergence of the state on the subsequence $y^{\alpha_i} \to y^* = Su^*$ in $H_0^1(\Omega) \cap C(\bar{\Omega})$. Now let $u \in U_{\mathrm{ad}}$ be arbitrary, then

$$J_0(u^*) = \lim_{i\to\infty} J_0(u^{\alpha_i}) = \lim_{i\to\infty} J_{\alpha_i}(u^{\alpha_i}) \leq \lim_{i\to\infty} J_{\alpha_i}(u) = J_0(u).$$

Hence u^* is a minimizer of J_0. The solution $\bar{u}$ of (P) is unique and since the problems (P) and (P^α) coincide for $\alpha = 0$ we obtain $\bar{u} = u^*$. As the norm is weakly lower semicontinuous we get

$$\limsup_{i\to\infty} \|u^{\alpha_i}\| \leq \|u^*\|_{L^2(\Omega)} \leq \liminf_{i\to\infty} \|u^{\alpha_i}\|_{L^2(\Omega)} \leq \limsup_{i\to\infty} \|u^{\alpha_i}\|_{L^2(\Omega)}$$

which shows $\|u^{\alpha_i}\|_{L^2(\Omega)} \to \|u^*\|_{L^2(\Omega)}$. As a well known fact, weak and norm convergence yield strong convergence and hence we have $u^{\alpha_i} \to u^*$. As the sequence u^{α_i} was arbitrarily chosen we obtain convergence of the whole sequence $u^\alpha \to \bar{u}$, which finishes the proof. $\qquad\square$

3.1 Optimality conditions

Let us assume that the Slater condition given in Assumption 2 is satisfied. Then first order necessary optimality conditions can be established for the regularized problem.

Theorem 3.2. *Let (u^α, y^α) be the solution of the problem (P^α). Furthermore, let Assumption 2 be fulfilled. Then, there exists an adjoint state $p^\alpha \in W^{1,s}(\Omega)$, $s \in [1, N/(N-1))$, a Lagrange multiplier $\mu^\alpha \in \mathcal{M}(\bar{\Omega})$ and a subdifferential $\lambda^\alpha \in \partial j(u^\alpha)$ such that the following optimality system holds:*

$$\begin{cases} Ay^\alpha = u^\alpha & in\ \Omega, \\ y^\alpha = 0 & on\ \Gamma, \end{cases} \tag{3a}$$

$$\begin{cases} A^*p^\alpha = y^\alpha - y_d + \mu^\alpha & in\ \Omega, \\ p^\alpha = 0 & on\ \Gamma, \end{cases} \tag{3b}$$

$$(p^\alpha + \alpha u^\alpha + \beta\lambda^\alpha, u - u^\alpha) \geq 0 \quad \forall u \in U_{ad}, \tag{3c}$$

$$\langle \mu^\alpha, y^\alpha - \psi \rangle_{\mathcal{M}(\bar{\Omega}), C(\bar{\Omega})} = 0, \quad \mu^\alpha \geq 0. \tag{3d}$$

Proof. The proof can be found in [10, Theorem 2.5]. $\qquad\square$

In the following we collect some results similar to Lemma 2.6.

Lemma 3.3. *Let $u^\alpha, y^\alpha, p^\alpha, \lambda^\alpha, \mu^\alpha$ satisfy the optimality system (3a)-(3d). Then the following relations hold:*

$$u^\alpha(x) = \begin{cases} u_a(x) & if \quad \beta - \alpha u_a(x) < p^\alpha(x) \\ \frac{1}{\alpha}(\beta - p^\alpha(x)) & if \quad \beta \leq p^\alpha(x) \leq \beta - \alpha u_a(x) \\ 0 & if \quad |p^\alpha(x)| < \beta \\ \frac{1}{\alpha}(-\beta - p^\alpha(x)) & if \quad -\alpha u_b(x) - \beta \leq p^\alpha(x) \leq -\beta \\ u_b(x) & if \quad p^\alpha(x) < -\alpha u_b(x) - \beta \end{cases}$$

$$\lambda^\alpha(x) = P_{[-1,1]}\left(-\frac{1}{\beta}(p^\alpha(x) + \alpha u^\alpha(x))\right)$$

$$u^\alpha(x) = P_{[u_a(x),u_b(x)]}\left(-\frac{1}{\alpha}(p^\alpha(x) + \beta\lambda^\alpha(x))\right)$$

Proof. These results can be proven by using a pointwise interpretation of the optimality condition (3c). $\qquad\square$

In the subsequent analysis we will need that the multipliers for the problem (P^α) are uniformly bounded for all $\alpha \geq 0$. Note that for $\alpha = 0$ the problem (P^α) reduces to problem (P). The boundedness of the multiplier can be expected from abstract theory [5], and we make use of the Slater condition to prove it.

Lemma 3.4. *Let $\alpha \geq 0$ and define the set*

$$M^\alpha := \{\mu^\alpha \in \mathcal{M}(\bar{\Omega}) : (u^\alpha, y^\alpha, p^\alpha, \lambda^\alpha, \mu^\alpha) \text{ satisfy } (3a) - (3d)\}.$$

of all multipliers associated with problem (P^α). Then the multipliers are uniformly bounded, i.e. there exists a constant $C > 0$ independent from α such that

$$\|\mu^\alpha\|_{\mathcal{M}(\bar{\Omega})} \leq C, \quad \forall \mu^\alpha \in M^\alpha \quad \forall \alpha \geq 0.$$

Proof. We follow the book of Tröltzsch [31] and consider our solution mapping $S : L^2(\Omega) \to H_0^1(\Omega) \cap C(\bar{\Omega})$. Then the dual operator is a mapping $S^* : \mathcal{M}(\bar{\Omega}) \to L^2(\Omega)$. Let $\alpha \geq 0$ be given, and u^α, y^α be the solution of (P^α) with an associated multiplier μ^α. We now use the Slater condition from Assumption 2 and compute for any $f \in C(\bar{\Omega})$ with $\|f\|_\infty = 1$:

$$\sigma\left|\int_\Omega f d\mu^\alpha\right| \leq \sigma\int_\Omega |f| d\mu^\alpha \leq \int_\Omega \sigma d\mu^\alpha \leq \int_\Omega (\psi - \hat{y})d\mu^\alpha$$

$$= \underbrace{\langle\mu^\alpha, \psi - y^\alpha\rangle_{\mathcal{M}(\bar{\Omega}),C(\bar{\Omega})}}_{=0 \text{ by } (3d)} + \langle\mu^\alpha, y^\alpha - \hat{y}\rangle_{\mathcal{M}(\bar{\Omega}),C(\bar{\Omega})}$$

$$= \langle\mu^\alpha, S(u^\alpha - \hat{u})\rangle_{\mathcal{M}(\bar{\Omega}),C(\bar{\Omega})}$$

$$= \int_\Omega (S^*\mu^\alpha)(u^\alpha - \hat{u})\,dx.$$

Now recall that the adjoint equation (3b) can be rewritten as

$$S^*\mu^\alpha = S^*(y_d - Su^\alpha) - p^\alpha.$$

Furthermore by assumption $u^\alpha \in U_{ad}$ and by Theorem 2.3 and 2.5 we obtain that u^α, y^α and p^α are uniformly bounded in $L^2(\Omega)$. This now yields

$$\sigma\|\mu^\alpha\|_{\mathcal{M}(\bar\Omega)} = \sigma \sup_{f\in C(\bar\Omega),\ \|f\|_\infty=1} \left|\int_\Omega f\, d\mu^\alpha\right|$$

$$\leq \int_\Omega (S^*\mu^\alpha)(u^\alpha - \hat u)\, dx$$

$$= \int_\Omega (S^*(y_d - Su^\alpha) - p^\alpha)(u^\alpha - \hat u)\, dx$$

$$\leq c\|u^\alpha - \hat u\|_{L^2(\Omega)}(\|y_d - y^\alpha\|_{L^2(\Omega)} + \|p^\alpha\|_{L^2(\Omega)})$$

$$\leq c.$$

Dividing the above inequality by $\sigma > 0$ finishes the proof.

$\square$

4   The Augmented Lagrange Method

In the following we want to solve the regularized Problem (P^α) for $\alpha \to 0$. For fixed α we follow the idea presented in [22] and replace the inequality constraint $y \leq \psi$ by an augmented penalization term. In that way we get rid of the measure and work instead with an approximation.

4.1   The augmented Lagrange optimal control problem

First let us introduce the penalty function P which we use to augment the state constraints. Let $\rho > 0$ be a given penalty parameter, and let $\mu \in L^2(\Omega)$ with $\mu \geq 0$ be a given approximation of the Lagrange multiplier. Now we define

$$P(y,\rho,\mu) := \frac{1}{2\rho}\int_\Omega \left((\mu + \rho(y - \psi))_+\right)^2 - \mu^2 \ dx. \tag{4}$$

Let now $\rho > 0$ and $\mu \in L^2(\Omega)$ be given. Then in each step of the augmented Lagrange method the following sub-problem has to be solved: Minimize

$$J_\rho^\alpha(y,u,\mu) := \frac{1}{2}\|y - y_d\|_{L^2(\Omega)}^2 + \beta\|u\|_{L^1(\Omega)} + \frac{\alpha}{2}\|u\|_{L^2(\Omega)}^2 + P(y,\rho,\mu) \tag{$P_{\alpha,\rho,\mu}$}$$

with $\alpha > 0$, subject to the state equation and the control constraints

$$y = Su, \quad u \in U_{ad}.$$

A solution of ($P_{\alpha,\rho,\mu}$) will be denoted by u_ρ^α with associated state y_ρ^α and adjoint state p_ρ^α. The next theorem shows that the sub-problem is unique solve-able.

Theorem 4.1 (Existence of solutions of the augmented Lagrange sub-problem). *For every $\rho > 0$, $\mu \in L^2(\Omega)$ with $\mu \geq 0$ the augmented Lagrange control problem ($P_{\alpha,\rho,\mu}$) admits a unique solution $u_\rho^\alpha \in U_{ad}$ with associated optimal state $y_\rho^\alpha \in Y$ and adjoint state p_ρ^α.*

Proof. Since U_{ad} is closed, bounded and convex and J_ρ^α is coercive, weakly lower semi-continuous and strictly convex, problem ($P_{\alpha,\rho,\mu}$) has a unique solution $u_\rho^\alpha \in U_{ad}$. For more details see [31] and [12]. $\square$

Theorem 4.2 (First-order necessary optimality conditions). *Let $(u_\rho^\alpha, y_\rho^\alpha)$ be the solution of ($P_{\alpha,\rho,\mu}$). Then, there exists a unique adjoint state $p_\rho^\alpha \in H_0^1(\Omega)$ associated with the optimal control u_ρ^α and a subdifferential $\lambda_\rho^\alpha \in \partial j(u_\rho^\alpha)$, satisfying the following system.*

$$\begin{cases} Ay_\rho^\alpha = u_\rho^\alpha & in\ \Omega, \\ y_\rho^\alpha = 0 & on\ \Gamma, \end{cases} \tag{5a}$$

$$\begin{cases} A^* p_\rho^\alpha = y_\rho^\alpha - y_d + \mu_\rho^\alpha & in\ \Omega, \\ p_\rho^\alpha = 0 & on\ \Gamma, \end{cases} \tag{5b}$$

$$(p_\rho^\alpha + \alpha u_\rho^\alpha + \beta \lambda_\rho^\alpha, u - u_\rho^\alpha) \geq 0 \quad \forall u \in U_{ad}, \tag{5c}$$

$$\mu_\rho^\alpha := \left(\mu + \rho(y_\rho^\alpha - \psi)\right)_+. \tag{5d}$$

Proof. Can be proven by extending the results in [18, Corollary 1.3, p. 73]. $\qquad\square$

Further we make an analogue observation like in [22]. Boundedness of μ_ρ^α in the L^1-norm is enough to get boundedness of the solution $(u_\rho^\alpha, y_\rho^\alpha, p_\rho^\alpha)$ of (5).

Theorem 4.3. *Let $\rho > 0$ and $\mu \in L^2(\Omega)$ be given. Let $s \in [1, N/(N-1))$ and α be bounded. Then there is a constant $c > 0$ independent of α, ρ, and μ such that for all solutions $(u_\rho^\alpha, y_\rho^\alpha, p_\rho^\alpha, \mu_\rho^\alpha)$ of (5) it holds*

$$\|y_\rho^\alpha\|_{H^1(\Omega)} + \|y_\rho^\alpha\|_{C(\bar\Omega)} + \|u_\rho^\alpha\|_{L^2(\Omega)} + \|p_\rho^\alpha\|_{W^{1,s}(\Omega)} \leq c(\|\mu_\rho^\alpha\|_{L^1(\Omega)} + 1).$$

Proof. The proof just differs from the one of [22, Theorem 3.3] concerning the additional sub-differential in (5b). Hence, we give just the most important steps here. Let us test the state equation (5a) with p_ρ^α and the adjoint equation (5b) with y_ρ^α. This yields

$$(p_\rho^\alpha, u_\rho^\alpha) = (y_\rho^\alpha - y_d, y_\rho^\alpha) + (\mu_\rho^\alpha, y_\rho^\alpha).$$

Now fix a $u \in U_{ad}$ and use it in (5c), yielding

$$(y_\rho^\alpha - y_d, y_\rho^\alpha) + (\mu_\rho^\alpha, y_\rho^\alpha) \leq (\alpha u_\rho^\alpha, u - u_\rho^\alpha) + (p_\rho^\alpha, u) + (\beta \lambda_\rho^\alpha, u - u_\rho^\alpha).$$

By Young's inequality and exploiting $(\lambda_\rho^\alpha, u - u_\rho^\alpha) \leq \|u\|_{L^1(\Omega)} - \|u_\rho^\alpha\|_{L^1(\Omega)}$, we have

$$\frac{1}{2}\left\|y_\rho^\alpha\right\|^2 + \frac{\alpha}{2}\left\|u_\rho^\alpha\right\|^2 \leq \frac{1}{2}\|y_d\|^2_{L^2(\Omega)} + \|\mu_\rho^\alpha\|_{L^1(\Omega)}\|y_\rho^\alpha\|_{C(\bar\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} + \|p_\rho^\alpha\|_{L^2(\Omega)}\|u\|_{L^2(\Omega)}$$
$$+ \beta \left(\|u\|_{L^1(\Omega)} - \|u_\rho^\alpha\|_{L^1(\Omega)}\right).$$

Let us fix $\bar s \in (1, N/(N-1))$ such that $W^{1,\bar s}(\Omega)$ is continuously embedded in $L^2(\Omega)$. From Theorem 2.1 we now get $\|y_\rho^\alpha\|_{H^1(\Omega)} + \|y_\rho^\alpha\|_{C(\bar\Omega)} \leq c\|u_\rho^\alpha\|_{L^2(\Omega)}$ and from Theorem 2.5 we get $\|p_\rho^\alpha\|_{L^2(\Omega)} \leq c\left(\|y_\rho^\alpha\|_{L^2(\Omega)} + \|y_d\|_{L^2(\Omega)} + \|\mu_\rho^\alpha\|_{L^1(\Omega)}\right)$. Now using the fact that u_ρ^α is bounded in $L^2(\Omega)$ and u is fixed to obtain the result.

$\qquad\square$

4.2 The general augmented Lagrange algorithm

In the following, let $(P^k_{\alpha,\rho,\mu})$ denote the augmented Lagrange sub-problem $(P_{\alpha,\rho,\mu})$ for given penalty parameter $\rho := \rho_k$, multiplier $\mu := \mu_k$ and regularization parameter $\alpha := \alpha_k$. We will denote its solution by $(\bar u_k, \bar y_k)$ with adjoint state $\bar p_k$ and updated multiplier $\bar\mu_k$, which is given by (5d).

Algorithm 1. *Let $\alpha_1 > 0$, $\rho_1 > 0$ and $\mu_1 \in L^2(\Omega)$ be given with $\mu_1 \geq 0$. Choose $\theta > 1$.*

1. *Solve $(P^k_{\alpha,\rho,\mu})$ and obtain $(\bar u_k, \bar y_k, \bar p_k)$.*

2. *Set $\bar\mu_k := (\mu_k + \rho_k(\bar y_k - \psi))_+$.*

3. *If the step is successful set $\mu_{k+1} := \mu_k$, $\rho_{k+1} := \rho_k$ and choose $0 < \alpha_{k+1}$ such that $\alpha_{k+1} < \alpha_k$.*

4. *Otherwise set $\mu_{k+1} := \bar\mu_k$ and $\alpha_{k+1} := \alpha_k$, increase penalty parameter $\rho_{k+1} := \theta\rho_k$.*

5. *If the stopping criterion is not satisfied set $k := k+1$ and go to step 1.*

Please note that we only decrease the regularization parameter α_k if the algorithm produces a successful step. Let us restate the system $(P^k_{\alpha,\rho,\mu})$ that is solved by $(\bar{u}_k, \bar{y}_k, \bar{p}_k, \bar{\mu}_k)$:

$$\begin{cases} A\bar{y}_k = \bar{u}_k & \text{in } \Omega, \\ \bar{y}_k = 0 & \text{on } \Gamma, \end{cases} \tag{6a}$$

$$\begin{cases} A^*\bar{p}_k = \bar{y}_k - y_d + \bar{\mu}_k & \text{in } \Omega, \\ \bar{p}_k = 0 & \text{on } \Gamma, \end{cases} \tag{6b}$$

$$\bar{u}_k \in U_{\text{ad}}, \tag{6c}$$

$$(\bar{p}_k + \alpha_k \bar{u}_k + \beta \bar{\lambda}_k, u - \bar{u}_k) \geq 0 \quad \forall u \in U_{\text{ad}}, \tag{6d}$$

$$\bar{\mu}_k := (\mu_k + \rho_k(\bar{y}_k - \psi))_+. \tag{6e}$$

4.3 The multiplier update rule

Let us start this section with a basic estimate, which will be useful in the sequel.

Lemma 4.4. *Let $\alpha_k > 0$ be given and let $(u^{\alpha_k}, y^{\alpha_k}, p^{\alpha_k}, \mu^{\alpha_k})$ be the solution of (3) and let $(\bar{u}_k, \bar{y}_k, \bar{p}_k, \bar{\mu}_k)$ solve (6). Then it holds*

$$\|y^{\alpha_k} - \bar{y}_k\|^2_{L^2(\Omega)} + \alpha_k \|u^{\alpha_k} - \bar{u}_k\|^2_{L^2(\Omega)} \leq (\bar{\mu}_k, \psi - \bar{y}_k) + \langle \mu^{\alpha_k}, \bar{y}_k - \psi \rangle. \tag{7}$$

Proof. Using (3c) and (6d), we obtain

$$\begin{aligned} 0 &\leq (p^{\alpha_k} - \bar{p}_k + \alpha_k(u^{\alpha_k} - \bar{u}_k) + \beta(\lambda^{\alpha_k} - \bar{\lambda}_k), \bar{u}_k - u^{\alpha_k}) \\ &= (S^*(Su^{\alpha_k} - S\bar{u}_k), \bar{u}_k - u^{\alpha_k}) - \alpha_k(\bar{u}_k - u^{\alpha_k}, \bar{u}_k - u^{\alpha_k}) \\ &\quad + (S^*(\mu^{\alpha_k} - \bar{\mu}_k), \bar{u}_k - u^{\alpha_k}) + \beta(\lambda^{\alpha_k} - \bar{\lambda}_k, \bar{u}_k - u^{\alpha_k}) \end{aligned} \tag{8}$$

Now we use that the subdifferential is a monotone operator, which yields $(\lambda^{\alpha_k} - \bar{\lambda}_k, \bar{u}_k - u^{\alpha_k}) \leq 0$. Note that $\lambda^{\alpha_k} \in \partial j(u^{\alpha_k})$ and $\bar{\lambda}_k \in \partial j(\bar{u}_k)$. This yields

$$\|y^{\alpha_k} - \bar{y}_k\|^2 + \alpha_k \|\bar{u}_k - u^{\alpha_k}\|^2 \leq (\bar{\mu}_k - \mu^{\alpha_k}, y^{\alpha_k} - \bar{y}_k)$$

The term on the right-hand side of equation (8) can be split into two parts:

$$(\bar{\mu}_k, y^{\alpha_k} - \bar{y}_k) = (\bar{\mu}_k, y^{\alpha_k} - \psi) + (\bar{\mu}_k, \psi - \bar{y}_k) \leq (\bar{\mu}_k, \psi - \bar{y}_k) \tag{9}$$

and

$$-\langle \mu^{\alpha_k}_k, y^{\alpha_k} - \bar{y}_k \rangle = -\langle \mu^{\alpha_k}, y^{\alpha_k} - \psi \rangle - \langle \mu^{\alpha_k}, \psi - \bar{y}_k \rangle = \langle \mu^{\alpha_k}, \bar{y}_k - \psi \rangle. \tag{10}$$

Here, we used the complementarity relation (3d) as well as $y^{\alpha_k} \leq \psi$ and $\bar{\mu}_k \geq 0$. Putting the inequalities (8), (9), and (10) together, we get

$$\|y^{\alpha_k} - \bar{y}_k\|^2_{L^2(\Omega)} + \alpha_k \|u^{\alpha_k} - \bar{u}_k\|^2_{L^2(\Omega)} \leq (\bar{\mu}_k, \psi - \bar{y}_k) + \langle \mu^{\alpha_k}, y_k - \psi \rangle.$$

which is the claim. $\qquad \square$

The following result motivates the update rule.

Lemma 4.5. *Let $(u^{\alpha_k}, y^{\alpha_k}, p^{\alpha_k}, \mu^{\alpha_k})$ and $(\bar{u}_k, \bar{y}_k, \bar{p}_k, \bar{\mu}_k)$ be given as in Lemma 4.4. Then it holds*

$$\frac{1}{\alpha_k} \|y^{\alpha_k} - \bar{y}_k\|^2_{L^2(\Omega)} + \|u^{\alpha_k} - \bar{u}_k\|^2_{L^2(\Omega)} \leq \frac{c}{\alpha_k} \left(\|(\bar{y}_k - \psi)_+\|_{C(\bar{\Omega})} + |(\bar{\mu}_k, \psi - \bar{y}_k)| \right). \tag{11}$$

Proof. From Lemma 4.4 we conclude using the estimate

$$\langle \mu^{\bar{\alpha}_k}, \bar{y}_k - \psi \rangle \leq \|\mu^{\alpha_k}\|_{\mathcal{M}(\Omega)} \|(\bar{y}_k - \psi)_+\|_{C(\bar{\Omega})}.$$

The result now follows using the uniform boundedness of μ^{α_k}, see Lemma 3.4. $\qquad \square$

This result shows that the iterates $(\bar{u}_k, \bar{y}_k)$ will converge to the solution of the regularized problem for fixed α_k if the quantity

$$\frac{1}{\alpha_k} \left(\|(\bar{y}_k - \psi)_+\|_{C(\bar{\Omega})} + |(\bar{\mu}_k, \psi - \bar{y}_k)| \right)$$

tends to zero for $k \to \infty$. To construct our update rule we follow the idea presented in [22] and define a step of Algorithm 1 to be successful if the condition

$$\frac{1}{\alpha_k} \left(\|(\bar{y}_k - \psi)_+\|_{C(\bar{\Omega})} + |(\bar{\mu}_k, \psi - \bar{y}_k)| \right) \leq \frac{\tau}{\alpha_n} \left(\|(\bar{y}_n - \psi)_+\|_{C(\bar{\Omega})} + |(\bar{\mu}_n, \psi - \bar{y}_n)| \right)$$

is satisfied with $\tau \in (0, 1)$. Here, we denoted by step n, $n < k$, the previous successful step. In [22] this quantity was also used as a stopping criterion. However this is not possible here, as we proceed to let α go to 0. Instead we will check the first order optimality conditions for problem (P) as a stopping criterion. This will be described in detail in section 6.

4.4 The augmented Lagrange algorithm in detail

Let us now formulate the algorithm based on the update rule established in the previous section.

Algorithm 2. *Let $\alpha_1 > 0$, $\rho_1 > 0$ and $\mu_1 \in L^2(\Omega)$ be given with $\mu_1 \geq 0$. Choose $\theta > 1, 0 < \omega < 1$, $\tau \in (0, 1)$ and $R_0^+ >> 1$. Set $k := 1$ and $n := 1$.*

1. *Solve $(P_{\alpha,\rho,\mu}^k)$ and obtain $(\bar{u}_k, \bar{y}_k, \bar{p}_k)$.*

2. *Set $\bar{\mu}_k := (\mu_k + \rho_k(\bar{y}_k - \psi))_+$.*

3. *Compute $R_k := \frac{1}{\alpha_k} \left(\|(\bar{y}_k - \psi)_+\|_{C(\bar{\Omega})} + |(\bar{\mu}_k, \psi - \bar{y}_k)| \right)$.*

4. *If $R_k \leq \tau R_{n-1}^+$ then the step k is successful, set*

$$\begin{cases} \alpha_{k+1} := \omega \alpha_k \\ \mu_{k+1} := \bar{\mu}_k \\ \rho_{k+1} := \rho_k \end{cases}$$

   and define $(u_n^+, y_n^+, p_n^+) := (\bar{u}_k, \bar{y}_k, \bar{p}_k)$, as well as $\mu_n^+ := \mu_{k+1}$ and $R_n^+ := R_k$. Set $n := n + 1$.

5. *Otherwise if the step k is not successful, set $\mu_{k+1} := \mu_k$ and $\alpha_{k+1} := \alpha_k$, and increase the penalty parameter $\rho_{k+1} := \theta \rho_k$.*

6. *If a stopping criterion is satisfied stop, otherwise set $k := k + 1$ and go to step 1.*

Again, please note that the regularization parameter α_k is only decreased when the algorithm produces a successful step. We will take advantage of this in the subsequent analysis.

4.5 Infinitely many successful steps

The main aim of this section is to prove that the proposed algorithm produces infinitely many successful steps. In order to prove this we consider the augmented Lagrange KKT system of the minimization problem

$$\text{Minimize } J_\rho^\alpha(y, u, \mu) = \frac{1}{2}\|y - y_d\|_{L^2(\Omega)}^2 + \beta\|u\|_{L^1(\Omega)} + \frac{\alpha}{2}\|u\|_{L^2(\Omega)}^2 + P(y, \rho, \mu)$$

subject to $y = Su$ and $u \in U_{\text{ad}}$. We fix the multiplier approximation μ, the regularization parameter α and let the penalization parameter ρ tend to infinity. As mentioned in [22] the problem reduces to a penalty method with additional shift parameter μ. The only difference to the approach in [22] is, that we have an additional L^1-term in the objective functional. However, taking a closer look at [22, Lemma 3.6] reveals that it also holds for an additional L^1-term. This yields the following Lemma.

Lemma 4.6. *Let $\mu \in L^2(\Omega)$ with $\mu \geq 0$ and $\alpha > 0$ be given. Let $(u_\alpha^\rho, y_\alpha^\rho, p_\alpha^\rho)$ be solutions of $(P_{\alpha,\rho,\mu})$ with $\rho > 0$ and (u^α, y^α) be the solution of (P^α). Then it holds $u_\rho^\alpha \to u^\alpha$ in $L^2(\Omega)$ and $y_\rho^\alpha \to y^\alpha$ in $H_0^1(\Omega) \cap C(\bar{\Omega})$ for $\rho \to \infty$.*

With a similar argument we can establish the next lemma. Again the proof can be found in [22, Lemma 3.7].

Lemma 4.7. *Under the same assumptions as in Lemma 4.6, it holds*

$$\lim_{\rho \to \infty} (\mu_\rho^\alpha, \psi - y_\rho^\alpha) = 0.$$

If we now combine these two results we can show that our algorithm produces infinitely many successful steps. This will be crucial in the convergence analysis in the next section.

Lemma 4.8. *The augmented Lagrange algorithm makes infinitely many successful steps.*

Proof. We assume that the algorithm produces only finitely many successful steps. Then there is an index m such that all steps $k > m$ are not successful. Due to the definition of the algorithm we obtain $\bar{\mu}_k = \bar{\mu}_m$ for all $k > m$ and $R_k > \tau R_m > 0$ as well as $\rho_k \to \infty$. This now yields a contradiction as with Lemma 4.6 and 4.7 we obtain

$$0 < \lim_{k \to \infty} R_k = \lim_{k \to \infty} \frac{1}{\alpha_k} \left(\left\| (\bar{y}_k - \psi)_+ \right\|_{C(\bar{\Omega})} + |(\bar{\mu}_k, \psi - \bar{y}_k)| \right) = 0$$

Please note that α_k is constant for $k > m$ since its value is only decreased in a successful step. $\qquad\square$

5 Convergence results

In this section we want to show convergence of Algorithm 2. Let us recall that the sequence (u_n^+, y_n^+, p_n^+) denotes the solution of the n-th successful iteration of Algorithm 2 with μ_n^+ being the corresponding approximation of the Lagrange multiplier. We start with proving L^1-boundedness of the Lagrange multipliers μ_n^+, which is accomplished in Lemma 5.2 below. To prove this result we need an auxiliary estimation first.

Lemma 5.1. *Let y_n^+, μ_n^+ be given as defined in Algorithm 2. Then it holds*

$$\frac{1}{\alpha_n} |(\mu_n^+, \psi - y_n^+)| \leq \frac{\tau^{n-1}}{\alpha_1} \left(\left\| (y_1^+ - \psi)_+ \right\|_{C(\bar{\Omega})} + \left\| \mu_1^+ \right\|_{L^2(\Omega)} \left\| (\psi - y_1^+)_+ \right\|_{L^2(\Omega)} \right). \tag{12}$$

Proof. Using the definition for a successful step we obtain:

$$\begin{aligned}
\frac{1}{\alpha_n} |(\mu_n^+, \psi - y_n^+)| &\leq \frac{\tau}{\alpha_{n-1}} \left(\left\| (y_{n-1}^+ - \psi)_+ \right\|_{C(\bar{\Omega})} + |(\mu_{n-1}^+, \psi - y_{n-1}^+)| \right) - \frac{1}{\alpha_n} \left\| (y_n^+ - \psi)_+ \right\|_{C(\bar{\Omega})} \\
&\leq \frac{\tau}{\alpha_{n-1}} \left\| (y_{n-1}^+ - \psi)_+ \right\|_{C(\bar{\Omega})} + \tau \left(\frac{1}{\alpha_{n-1}} |(\mu_{n-1}^+, \psi - y_{n-1}^+)| \right) \\
&\leq \frac{\tau}{\alpha_{n-1}} \left\| (y_{n-1}^+ - \psi)_+ \right\|_{C(\bar{\Omega})} \\
&\quad + \tau \left(\frac{\tau}{\alpha_{n-2}} \left(\left\| (y_{n-2}^+ - \psi)_+ \right\|_{C(\bar{\Omega})} + |(\mu_{n-2}^+, \psi - y_{n-2}^+)| \right) \right. \\
&\quad \left. - \frac{1}{\alpha_{n-1}} \left\| (y_{n-1}^+ - \psi)_+ \right\|_{C(\bar{\Omega})} \right) \\
&\leq \frac{\tau^2}{\alpha_{n-2}} \left(\left\| (y_{n-2}^+ - \psi)_+ \right\|_{C(\bar{\Omega})} + |(\mu_{n-2}^+, \psi - y_{n-2}^+)| \right)
\end{aligned}$$

The rest now follows by induction and a standard estimate.

$\qquad\square$

We want to point out that the right hand side of (12) goes to 0 as $n \to \infty$. This will be crucial in the following convergence analysis and is a result of our update rule. Let us now show the L^1-boundedness of the Lagrange multipliers (μ_n^+).

Lemma 5.2 (Boundedness of the Lagrange multiplier). *Let Assumption 2 be fulfilled. Then Algorithm 2 generates an infinite sequence of bounded iterates, i.e., there is a constant $C > 0$ such that for all n it holds*

$$\left\| y_n^+ \right\|_{H^1(\Omega)} + \left\| y_n^+ \right\|_{C(\bar{\Omega})} + \left\| u_n^+ \right\|_{L^2(\Omega)} + \left\| p_n^+ \right\|_{W^{1,s}(\Omega)} + \left\| \mu_n^+ \right\|_{L^1(\Omega)} \leq C.$$

Proof. Let $(\hat{u}, \hat{y})$ be the Slater point given by Assumption 2, i.e., there exists $\sigma > 0$, such that $\hat{y} + \sigma \leq \psi$. Then we can estimate

$$\sigma \|\mu_n^+\|_{L^1(\Omega)} = \int_\Omega \sigma \mu_n^+ \, dx \leq \int_\Omega \mu_n^+ (\psi - \hat{y}) \, dx = \int_\Omega \mu_n^+ (\psi - y_n^+ + y_n^+ - \hat{y}) \, dx$$

$$= \underbrace{\int_\Omega \mu_n^+ (\psi - y_n^+) \, dx}_{(I)} + \underbrace{\int_\Omega \mu_n^+ (y_n^+ - \hat{y}) \, dx}_{(II)}.$$

The first part (I) can be estimated with Lemma 5.1 yielding

$$(I) \leq |(\mu_n^+, \psi - y_n^+)| \leq \frac{\alpha_n}{\alpha_{n-1}} \tau^{n-1} \left(\left\| (y_1^+ - \psi)_+ \right\|_{C(\bar{\Omega})} + \left\| \mu_1^+ \right\|_{L^2(\Omega)} \left\| (\psi - y_1^+)_+ \right\|_{L^2(\Omega)} \right)$$
$$\leq c\tau^{n-1}. \tag{13}$$

Please note that we used the monotonicity of $(\alpha_n)_n$. Before we estimate part (II) recall that we have the inequality

$$(\lambda_n^+, u - u_n^+) \leq \|u\|_{L^1(\Omega)} - \|u_n^+\|_{L^1(\Omega)}$$

for every $u \in L^1(\Omega)$. By definition we obtain that $u \in U_{ad}$ implies $u \in L^\infty(\Omega)$. Now the second part (II) can be estimated using Young's Inequality as follows

$$\int_\Omega \mu_n^+ (y_n^+ - \hat{y}) \, dx = \langle A^* p_n^+ - (y_n^+ - y_d), y_n^+ - \hat{y} \rangle$$
$$= \langle p_n^+, A(y_n^+ - \hat{y}) \rangle - (y_n^+ - y_d, y_n^+ - \hat{y})$$
$$= (p_n^+, u_n^+ - \hat{u}) - (y_n^+ - y_d, y_n^+ - \hat{y})$$
$$\leq -(\alpha u_n^+, u_n^+ - \hat{u}) - (y_n^+ - y_d, y_n^+ - \hat{y}) - \beta(\lambda_n^+, u_n^+ - \hat{u})$$
$$\leq (\alpha u_n^+, \hat{u} - u_n^+) + (y_n^+ - y_d, \hat{y} - y_n^+) + \beta(\|\hat{u}\|_{L^1(\Omega)} - \|u_n^+\|_{L^1(\Omega)}) \tag{14}$$
$$= \alpha(u_n^+ - \hat{u}, \hat{u} - u_n^+) + \alpha(\hat{u}, \hat{u} - u_n^+) + (y_n^+ - \hat{y}, \hat{y} - y_n^+) + (\hat{y} - y_d, \hat{y} - y_n^+)$$
$$\quad + \beta(\|\hat{u}\|_{L^1(\Omega)} - \|u_n^+\|_{L^1(\Omega)})$$
$$\leq -\frac{\alpha}{2} \left\| \hat{u} - u_n^+ \right\|_{L^2(\Omega)}^2 - \frac{1}{2} \left\| \hat{y} - y_n^+ \right\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \left\| \hat{u} \right\|_{L^2(\Omega)}^2 + \frac{1}{2} \left\| \hat{y} - y_d \right\|_{L^2(\Omega)}^2$$
$$\quad + \beta(\|\hat{u}\|_{L^1(\Omega)} - \|u_n^+\|_{L^1(\Omega)}).$$

Putting (13) and (14) together yields

$$\left\| \mu_n^+ \right\|_{L^1(\Omega)} + \frac{\alpha}{2} \left\| \hat{u} - u_n^+ \right\|_{L^2(\Omega)}^2 + \frac{1}{2} \left\| \hat{y} - y_n^+ \right\|_{L^2(\Omega)}^2$$
$$\leq \frac{\tau^{n-1}}{\sigma} C + \frac{\alpha}{2} \|\hat{u}\|_{L^2(\Omega)}^2 + \beta \|\hat{u}\|_{L^1(\Omega)} + \frac{1}{2} \|\hat{y} - y_d\|_{L^2(\Omega)}^2.$$

Since $\tau \in (0, 1)$ by assumption, the right-hand side is bounded. Consequently we get boundedness of (u_n^+) in $L^2(\Omega)$ and boundedness of (μ_n^+) in $L^1(\Omega)$. By the regularity result Theorem 2.1, the sequence (y_n^+) is uniformly bounded in $H_0^1(\Omega) \cap C(\bar{\Omega})$. Boundedness of (p_n^+) follows directly from Theorem 4.3. $\qquad \square$

Theorem 5.3 (**Convergence of solutions of the augmented Lagrange algorithm**). *As* $n \to \infty$ *we have for the sequence* (u_n^+, y_n^+) *generated by Algorithm 2*

$$(u_n^+, y_n^+) \to (\bar{u}, \bar{y}), \ in \ L^2(\Omega) \times (H_0^1(\Omega) \cap C(\bar{\Omega})).$$

where $\bar{u}$ *denotes the unique minimum norm solution of problem* (P).

Proof. Since the algorithm yields an infinite number of successful steps (Lemma 4.8) we get

$$\lim_{n \to \infty} R_n^+ = \lim_{n \to \infty} \frac{1}{\alpha_n} \left(\left\| (y_n^+ - \psi)_+ \right\|_{C(\bar{\Omega})} + |(\mu_n^+, \psi - y_n^+)| \right) = 0. \tag{15}$$

with $\alpha_n \to 0$. Let $(u^{\alpha_n}, y^{\alpha_n}, p^{\alpha_n}, \mu^{\alpha_n})$ be a solution of (3) for $\alpha := \alpha_n$ then we obtain from Lemma 4.4 the following inequality

$$\frac{1}{\alpha_n} \left\| y^{\alpha_n} - y_n^+ \right\|_{L^2(\Omega)}^2 + \left\| u^{\alpha_n} - u_n^+ \right\|_{L^2(\Omega)}^2 \leq \frac{1}{\alpha_n} \left(\langle \mu^{\alpha_n}, y_n^+ - \psi \rangle + |(\mu_n^+, \psi - y_n^+)| \right)$$

$$\leq \frac{1}{\alpha_n} \left(\left\| \mu^{\alpha_n} \right\|_{\mathcal{M}(\bar{\Omega})} \left\| (y_n^+ - \psi)_+ \right\|_{C(\bar{\Omega})} + |(\mu_n^+, \psi - y_n^+)| \right)$$

$$\leq \frac{c}{\alpha_n} \left(\left\| (y_n^+ - \psi)_+ \right\|_{C(\bar{\Omega})} + |(\mu_n^+, \psi - y_n^+)| \right).$$

Note that in the last step we used Lemma 3.4. With (15) from above, we conclude

$$\lim_{n \to \infty} \frac{1}{\alpha_n} \left\| y^{\alpha_n} - y_n^+ \right\|_{L^2(\Omega)}^2 + \left\| u^{\alpha_n} - u_n^+ \right\|_{L^2(\Omega)}^2 = 0.$$

As $\alpha_n \to 0$ with $n \to \infty$ we obtain by Lemma 3.1 that $u^{\alpha_n} \to \bar{u}$. Triangular inequality now reveals

$$\|u_n^+ - \bar{u}\|_{L^2(\Omega)} \leq \|u_n^+ - u^{\alpha_n}\|_{L^2(\Omega)} + \|u^{\alpha_n} - \bar{u}\|_{L^2(\Omega)} \to 0.$$

Convergence of $y_n^* \to \bar{y}$ follows from Theorem 2.1 which finishes the proof. □

Remark 1. *For the sequence* (y_n^+) *generated by Algorithm 2 we obtain*

$$\frac{1}{\alpha_n} \|y_n^+ - \bar{y}\|_{L^2(\Omega)}^2 \to 0$$

which is similar to the results obtained for a Tikhonov regularization without state constraints, see [32].

Let us assume that the adjoint state $\bar{p}$ and the multiplier corresponding to the state constraint $\bar{\mu}$ are unique then following [22, Theorem 3.12, Corollary 3.12] we get the following convergence result.

Theorem 5.4. *Let* $(\bar{u}, \bar{y}, \bar{p}, \bar{u})$ *satisfy the KKT-system* (1)*. Let us assume that* $(\bar{p}, \bar{\mu})$ *are uniquely given. Then it holds*

$$p_n^+ \rightharpoonup \bar{p} \qquad in \ L^2(\Omega),$$
$$\mu_n^+ \overset{*}{\rightharpoonup} \bar{\mu} \qquad in \ \mathcal{M}(\bar{\Omega}).$$

6 Numerical method in detail

All optimal control problems have been solved using the above stated augmented Lagrange algorithm implemented with FEniCS [23] using the DOLFIN [24] Python interface. The arising sub-problems $(P_{\alpha,\rho,\mu})$ have been solved combining two methods. The first method is the active set method presented by Stadler [30], where optimal control problems of type $(P_{\alpha,\rho,\mu})$ have been solved, but without augmented state constraints. The second is the method established by Ito et al. [20] that presented an active set method for optimal control problems with state constraints but without a L^1-cost term.

Like in [30] we set for (6)

$$\bar{\xi}_k = \lambda_k^\alpha - \lambda_k^a + \lambda_k^b,$$

where λ_k^α denotes the subdifferential of $\beta\|\bar{u}_k\|_{L^1(\Omega)}$, λ_k^a the multiplier to the lower control constraints $u_a - \bar{u}_k \leq 0$ and λ_k^b the multiplier corresponding to the upper control constraint $\bar{u}_k - u_b \leq 0$. Then (5b) can be written as

$$\bar{p}_k + \alpha\bar{u}_k + \bar{\xi}_k = 0.$$

Defining the following active sets, see also Lemma 3.3:

$$\begin{aligned}
\mathcal{Y}_-^k &= \{x \in \Omega : (\mu + \rho(y_{k+1} - \psi)) > 0\}, \\
\mathcal{Y}_+^k &= \Omega \setminus \mathcal{Y}_-^k, \\
\mathcal{A}_a^k &= \{x \in \Omega : p_k > \beta - \alpha u_a\}, \\
\mathcal{A}_0^k &= \{x \in \Omega : |p_k| < \beta\}, \\
\mathcal{A}_b^k &= \{x \in \Omega : p_k < -\alpha u_b - \beta\}, \\
\mathcal{I}_-^k &= \{x \in \Omega : \beta \leq p_k \leq \beta - \alpha u_a\}, \\
\mathcal{I}_+^k &= \{x \in \Omega : -\alpha u_b - \beta \leq p_k \leq -\beta\},
\end{aligned}$$

The resulting sub-problem of the augmented Lagrange method can now be solved by the following algorithm.

Active set method (inner iteration)

1. Choose initial data u_0, p_0 and parameters α, ρ, compute the active sets $\mathcal{Y}_-^0$, $\mathcal{Y}_+^0$, $\mathcal{A}_a^0$, $\mathcal{A}_0^0$, $\mathcal{A}_b^0$, $\mathcal{I}_-^0, \mathcal{I}_+^0$ and set $k := 0$.

2. Solve for $(u_{k+1}, y_{k+1}, p_{k+1}, \xi_{k+1})$ satisfying

$$\begin{aligned}
Ay_{k+1} - u_{k+1} - f &= 0, \\
-A^* p_{k+1} + y_{k+1} - y_d + \mu_{k+1} &= 0, \\
p_{k+1} + \alpha u_{k+1} + \xi_{k+1} &= 0,
\end{aligned} \tag{16}$$

$$u_{k+1} = \begin{cases} u_a & \text{on } \mathcal{A}_a^k, \\ 0 & \text{on } \mathcal{A}_0^k, \\ u_b & \text{on } \mathcal{A}_b^k, \end{cases} \qquad \xi_{k+1} = \begin{cases} -\beta & \text{on } \mathcal{I}_-^k, \\ \beta & \text{on } \mathcal{I}_+^k, \end{cases} \qquad \mu_{k+1} = \begin{cases} 0 & \text{on } \mathcal{Y}_-^k, \\ \mu + \rho(y_{k+1} - \psi) & \text{on } \mathcal{Y}_+^k. \end{cases} \tag{17}$$

3. Compute the active sets $\mathcal{Y}_-^{k+1}, \mathcal{Y}_+^{k+1}, \mathcal{A}_a^{k+1}, \mathcal{A}_0^{k+1}, \mathcal{A}_b^{k+1}, \mathcal{I}_-^{k+1}, \mathcal{I}_+^{k+1}$

4. If the following equalities hold: $\mathcal{A}_a^{k+1} = \mathcal{A}_a^k$, $\mathcal{A}_0^{k+1} = \mathcal{A}_0^k$, $\mathcal{A}_b^{k+1} = \mathcal{A}_b^k$, $\mathcal{I}_-^{k+1} = \mathcal{I}_-^k$, $\mathcal{I}_+^{k+1} = \mathcal{I}_+^k$, $\mathcal{Y}_-^{k+1} = \mathcal{Y}_-^k$ and $\mathcal{Y}_+^{k+1} = \mathcal{Y}_+^k$ then go step 5. Otherwise set $k = k + 1$ and go to step 2.

5. Compute the subdifferential $\lambda_{k+1} := P_{[-1,1]}\left(-\frac{1}{\beta}\xi_{k+1}\right)$.

The computation of the L^1-subdifferential follows from a projection formular similar to the one from Lemma 3.3. Since the active sets are disjoint subsets of Ω the calculation of ξ_{k+1} in Step 4 does not evoke any conflicts to its usage on the subsets $\mathcal{I}_-^k, \mathcal{I}_+^k$ in Step 3 of the algorithm. Further, the termination criterion yields a solution of the augmented Lagrange subproblem (P^α).

Lemma 6.1. *If $\mathcal{A}_a^{k+1} = \mathcal{A}_a^k, \mathcal{A}_0^{k+1} = \mathcal{A}_0^k, \mathcal{A}_b^{k+1} = \mathcal{A}_b^k, \mathcal{I}_-^{k+1} = \mathcal{I}_-^k, \mathcal{I}_+^{k+1} = \mathcal{I}_+^k, \mathcal{Y}_-^{k+1} = \mathcal{Y}_-^k, \mathcal{Y}_+^{k+1} = \mathcal{Y}_+^k$, then $(u_k, y_k, p_k, \mu_k, \lambda_k)$ is the solution $(\bar{u}_k, \bar{y}_k, \bar{p}_k, \bar{\mu}_k, \bar{\lambda}_k)$ to (6) with α, μ, β fixed.*

Proof. Since for given active sets the solution to (16) is unique we have $(u_{k+1}, y_{k+1}, p_{k+1}) = (u_k, y_k, p_k)$. By definition of the active sets $\mathcal{Y}_-^k, \mathcal{Y}_+^k$ we get $\mu_{k+1} = (\mu + \rho(y_{k+1} - \psi))_+$. The optimality condition (6d) can be equivalently expressed by

$$\bar{u}_k - \max(0, \bar{u}_k + c(\bar{\xi}_k - \beta)) - \min(0, \bar{u}_k + c(\bar{\xi}_k + \beta))$$
$$+ \max(0, (\bar{u}_k - u_b) + c(\bar{\xi}_k - \beta)) + \min(0, (\bar{u}_k - u_a) + c(\bar{\xi}_k + \beta)) = 0.$$

where $c > 0$ arbitrary. Choosing $c = \alpha^{-1}$ and exploiting $\xi_\alpha^\rho = -(p_\rho^\alpha + \alpha \bar{u}_k)$ we get

$$\bar{u}_k - \alpha^{-1} \max(0, -\bar{p}_k - \beta) - \alpha^{-1} \min(0, -\bar{p}_k + \beta)$$
$$+ \alpha^{-1} \max(0, -\bar{p}_k - \beta - \alpha u_b) + \alpha^{-1} \min(0, -\bar{p}_k + \beta - \alpha u_a) = 0,$$

which is satisfied for $\bar{u}_k = u_{k+1}, \bar{p}_k = p_{k+1}$ defined by (17). Moreover, λ_{k+1} satisfies (6b) by definition. Consequently $(u_k, y_k, p_k, \mu_k, \lambda_k)$ satisfies (6). $\qquad\square$

However, high values of the penalty parameter ρ paired with small values of the Tikhonov parameter α may evoke bad stability during solution of the subproblem. To counteract this aspect we introduce a so called intermediate step. Here, Step 3 and Step 4 of Algorithm 2 are extended for a third alternative. If the current iterates of the k-th iteration do not satisfy the update rule but sufficiently satisfy the feasibility and complementarity condition, i.e.

$$R_k \geq \tau R_{n-1}^+ \quad \text{and} \quad \|(\bar{y}_k - \psi)_+\|_{C(\bar{\Omega})} + |(\bar{\mu}_k, \psi - \bar{y}_k)| < \varepsilon_I,$$

with $\varepsilon_I > 0$, we set

$$\alpha_{k+1} = \omega \alpha_k,$$
$$\mu_{k+1} = \bar{\mu}_k,$$
$$(u_n^+, y_n^+, p_n^+) := (\bar{u}_k, \bar{y}_k, \bar{p}_k).$$

As a termination criterion we check the optimality conditions of the current iterate $(u_n^+, y_n^+, p_n^+, \mu_n^+, \lambda_n^+)$ i.e. we stop the algorithm if the inequality

$$\left\| u_n^+(x) - P_{[u_a(x), u_b(x)]}\left(u_n^+(x) - (p_n^+ + \beta\lambda_n^+(x)))\right)\right\|_{L^2(\Omega)}$$
$$+ \left\|(y_n^+(x) - \psi)_+\right\|_{C(\bar{\Omega})} + |(\mu_n^+(x), y_n^+(x) - \psi)| \leq \varepsilon \tag{18}$$

is satisfied. In order to be consistent we set $\varepsilon_I < \varepsilon$.

As the Active-Set methods are related to the class of semi-smooth Newton methods we cannot expect a global convergence behavior of the method described above. Furthermore, the problem becomes bad conditioned if $\alpha \to 0$ or $\rho \to \infty$. Due to the intermediate step we expect ρ to be bounded. However as α goes to zero we have to globalize our method. We use a projected gradient method to construct suitable starting values for the Active-Set method.

7 Numerical results

Let us present some numerical results to support our method. We apply our method for problems of the following form:

$$\min \ J(y, u) := \frac{1}{2}\|y - y_d\|_{L^2(\Omega)}^2 + \beta\|u\|_{L^1(\Omega)} \tag{19}$$

subject to

$$
\begin{aligned}
Ay &= u + f && \text{in } \Omega, \\
y &= 0 && \text{on } \partial\Omega, \\
y &\leq \psi && \text{in } \Omega, \\
u_a \leq u &\leq u_b && \text{in } \Omega.
\end{aligned}
$$

The additional variable $f \in L^2(\Omega)$ allows us to construct test problems with known solutions.

7.1 Example 1: Bang-Bang-Off Example in One Space Dimension

We first consider the one-dimensional case and define $\Omega = (-1, 1)$, $u_a = -1$, $u_b = 1$ and $\beta = 1$. Furthermore set

$$
\bar{y}(x) := \begin{cases}
28 + 108 \cdot x + 144 \cdot x^2 + 64 \cdot x^3 & \text{if } x \in [-1, -\frac{3}{4}] \\
1 & \text{if } x \in [-\frac{3}{4}, \frac{3}{4}] \\
28 - 108 \cdot x + 144 \cdot x^2 - 64 \cdot x^3 & \text{if } x \in [\frac{3}{4}, 1]
\end{cases}
$$

$$
\bar{p}(x) := -2\cos\left(\frac{3\pi}{2}x\right)
$$

$$
\bar{u}(x) := \begin{cases}
0 & \text{if } x \in [-1, -\frac{8}{9}] \cup [-\frac{4}{9}, -\frac{2}{9}] \cup [\frac{2}{9}, \frac{4}{9}] \cup [\frac{8}{9}, 1] \\
1 & \text{if } x \in (-\frac{2}{9}, \frac{2}{9}) \\
-1 & \text{if } x \in (-\frac{8}{9}, -\frac{4}{9}) \cup (\frac{4}{9}, \frac{8}{9})
\end{cases}
$$

$$
\bar{\mu}(x) := \begin{cases}
\mathrm{Exp}\left(-\frac{1}{1-\left(\frac{4}{3}x\right)^2}\right) & \text{if } x \in [-\frac{3}{4}, \frac{3}{4}] \\
0 & \text{else}
\end{cases}
$$

$$
\psi(x) := 1.
$$

Some calculations show that $\bar{y}, \bar{p} \in C^2(\Omega)$ and $\bar{y} = \bar{p} = 0$ on $\partial\Omega$. By construction we obtain $\bar{u}(x) \in \{-1, 0, 1\}$ for a.e. $x \in \Omega$. In order to satisfy the optimality conditions we now set

$$
\begin{aligned}
f(x) &:= -\Delta\bar{y}(x) - \bar{u}(x), \\
y_d(x) &:= \Delta\bar{p}(x) + \bar{y}(x) + \bar{\mu}(x).
\end{aligned}
$$

One now can check that the functions $(\bar{u}, \bar{y}, \bar{p}, \bar{\mu})$ satisfy the KKT conditions defined in Theorem 2.4 with a suitable modification for the forward equation. We apply our algorithm with the following set of parameters

$$
\theta = 5, \quad \omega = 0.75, \quad \tau = 0.8, \quad \varepsilon = 10^{-6}, \quad \varepsilon_I = 5 \cdot 10^{-7}.
$$

The interval Ω is divided into 10^6 equidistant elements. The algorithm stops after a total of 40 iterations, which splits in 13 successful, 19 intermediate and 8 not successful iterations with an average of 5.25 inner iterations. The parameters were initialized with $\alpha := 1$ and $\rho := 100$ and the final parameters are $\alpha = 0.75^{32} \approx 10^{-4}$ and $\rho = 100 \cdot 5^8 \approx 3.9 \cdot 10^7$.
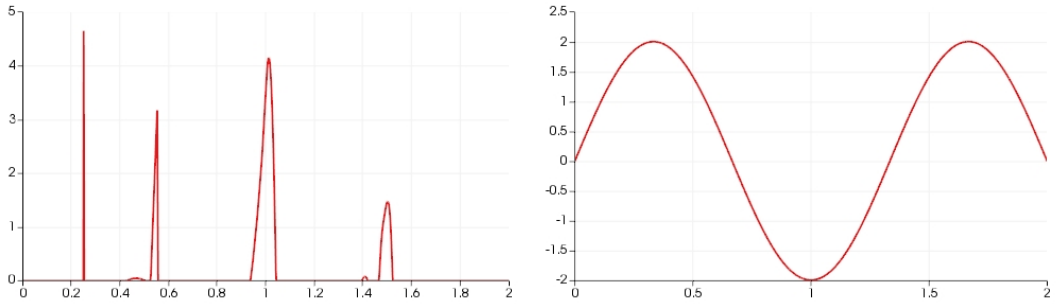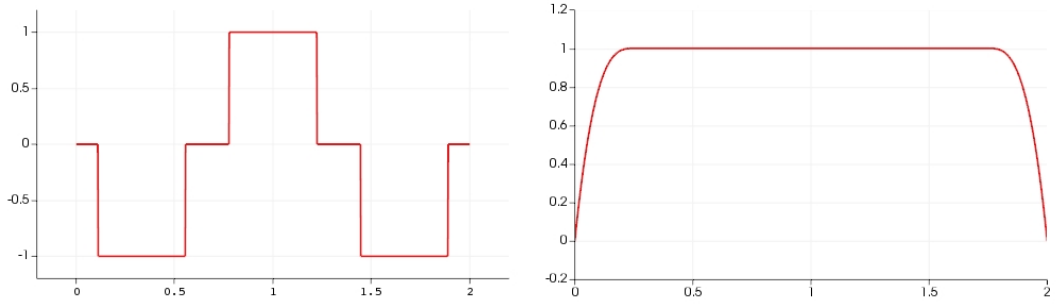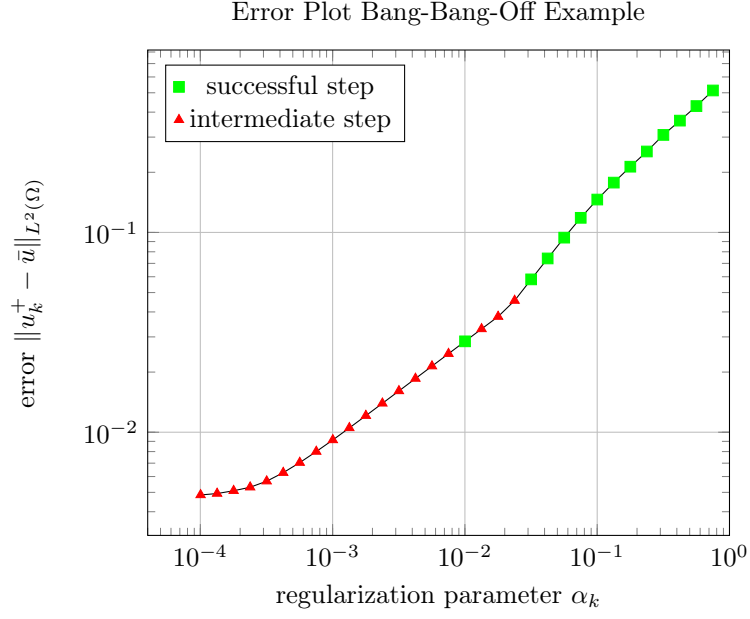
As we have an exact solution we can compute convergence rates. We plot the L^2-error $\|u_k^+ - \bar{u}\|_{L^2(\Omega)}$ over the regularization parameter α_k. Note the we only plot successful and intermediate steps. As expected we see that the algorithm produces only intermediate steps after some given time. The error can be found in Figure 1 and plots of the computed solution can be seen in Figure 2 and 3.

Remark 2. *Analysing the error $\|u_k^+ - \bar{u}\|_{L^2(\Omega)}$ we see that the error behaves like*

$$
\|u_k^+ - \bar{u}\|_{L^2(\Omega)} = \mathcal{O}\left(\alpha^\gamma\right) \tag{20}
$$

with constant $\alpha, \gamma > 0$. We want to mention that the exact control $\bar{u}$ satisfies the following regularity assumption

$$
meas\{x \in \Omega : \left| |\bar{p}(x)| - \beta \right| < \varepsilon\} \leq c\varepsilon^\kappa,
$$

17

Error Plot Bang-Bang-Off Example

Figure 1: Error $\|u_k^+ - \bar{u}\|_{L^2(\Omega)}$ over α_k for example 1.

Figure 2: Computed control u and state y for example 1.

Figure 3: Computed multiplier μ and adjoint state p for example 1.

18

holds for all $\varepsilon > 0$ with some $\kappa > 0$, which can be used to prove error estimates of the form (20) for some algorithms, see e.g. [27, 34]. However it is an open problem to prove convergence rates for the augmented Lagrange method presented in this paper.

7.2 Example 2: Bang-Bang-Off Example in Two Space Dimension

We set $u_a = -1$, $u_b = 1$. Let Ω be the circle around 0 with radius 2. We now define the following functions. For clarity and to shorten our notation we set $r := r(x, y) := \sqrt{x^2 + y^2}$.

$$\bar{y}(x, y) := \begin{cases} 1 & \text{if } r < 1 \\ 32 - 120 \cdot r + 180 \cdot r^2 - 130 \cdot r^3 + 45 \cdot r^4 - 6 \cdot r^5 & \text{if } r \geq 1 \end{cases}$$

$$\bar{p}(x, y) := \sin(x) \cdot \sin(y) \cdot \left(1 - \frac{5}{4}r^3 + \frac{15}{16}r^4 - \frac{3}{16}r^5\right)$$

$$\bar{u}(x, y) := -\text{Sign}(p(x, y))$$

$$\bar{\mu}(x, y) := \begin{cases} \text{Exp}\left(-\frac{1}{1-r^2}\right) & \text{if } r < 1 \\ 0 & \text{if } r \geq 1 \end{cases}$$

$$\psi(x, y) := 1$$

Some calculation show that $\bar{\mu}, \bar{p} \in C^2(\bar{\Omega})$ and $\bar{\mu} \in C(\bar{\Omega})$. Furthermore $\bar{y} = \bar{p} = 0$ on $\partial\Omega$. We now set

$$f(x, y) := -\Delta\bar{y}(x, y) - \bar{u}(x, y),$$
$$y_d(x, y) := \Delta\bar{p}(x, y) + \bar{y}(x, y) + \bar{\mu}(x, y).$$

One now can check that for $\beta = 0$ the functions $(\bar{u}, \bar{y}, \bar{p}, \bar{\mu})$ satisfy the KKT conditions defined in Theorem 2.4 leading to a bang-bang solution. For $\beta \neq 0$ we expect the optimal solution to omit a bang-bang-off structure. Here no exact solution is known. We computed this problem for different values of β on a regular triangular grid with approximately $1.8 \cdot 10^5$ degrees of freedom. The parameter used for this computation are $\tau = 0.8$, $\omega = 0.75$, $\theta = 5$, $\varepsilon = 10^{-6}$ and $\varepsilon_I = 5 \cdot 10^{-7}$. We started with $\alpha = 0.1$ and $\rho = 100$. Additional information for the calculations can be found in Table 1 while the computed controls can be seen in Figure 4.

β	final α	final ρ	successful steps	intermediate steps	not success-ful steps	average inner iterations
0.05	$2.38 \cdot 10^{-5}$	10^9	15	14	7	2.9
0.1	$2.38 \cdot 10^{-5}$	10^9	16	13	7	3.1
0.2	$3.17 \cdot 10^{-5}$	10^9	18	10	7	3.0
1	$5.6 \cdot 10^{-5}$	10^{10}	20	6	8	3.5

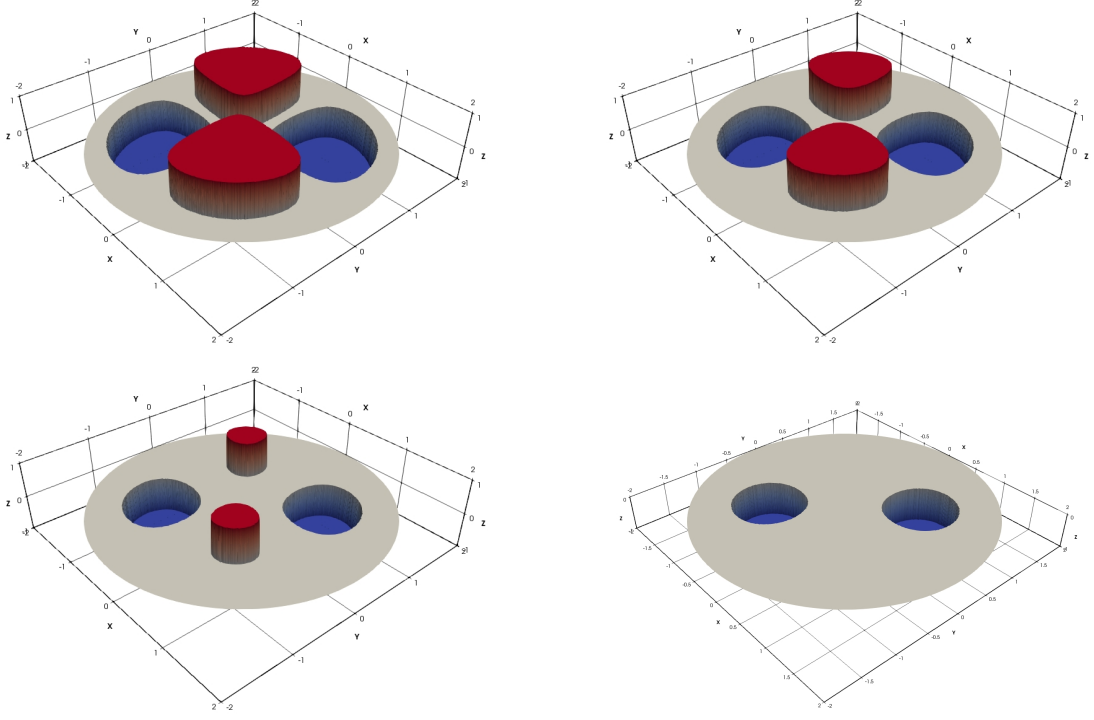Table 1: Additional information for the computation of example 2 for different β.

Figure 4: Computed discrete control for example 2 for different values of β. From left to right and from top to bottom: $\beta = 0.05$, $\beta = 0.1$, $\beta = 0.2$, $\beta = 1$.

7.3 Example 3

For the next example we set $\Omega = (0, 1)^2$, $u_a = -1$, $u_b = 1$ and $\beta = 10^{-3}$. Furthermore $\tau = 0.8$, $\omega = 0.75$ and $\theta = 10$. Now define

$$\psi(x, y) := 0.01$$

$$y_d(x, y) := \frac{1}{2\pi} \sin(\pi x) \sin(\pi y)$$

Note that here no exact solution is available. If the state constraints are neglected the exact solution is given by

$$\bar{y}(x, y) := y_d(x, y)$$
$$\bar{u}(x, y) := \Delta y_d(x, y).$$

This example is taken from [26] and is an example of an optimal control problem where the desired state is reachable and the source condition $\bar{u} = S^* w$ with an element $w \in L^2(\Omega)$ is satisfied if the state constraints are not present. We computed the solution on a regular triangular grid with $1.6 \cdot 10^5$ degrees of freedom, $\varepsilon = 10^{-6}$ and $\varepsilon_I = 5 \cdot 10^{-7}$. As starting values we set $\alpha = 0.1$ and $\rho = 100$. The algorithm stopped after 8 successful, 25 intermediate and 9 not successful steps with the final values $\alpha = 0.1 \cdot 0.75^{33} \approx 7.5 \cdot 10^{-6}$ and $\rho = 100 \cdot 5^9 \approx 2.0 \cdot 10^8$. The computed results can be seen in Figure 5 and Figure 6 .
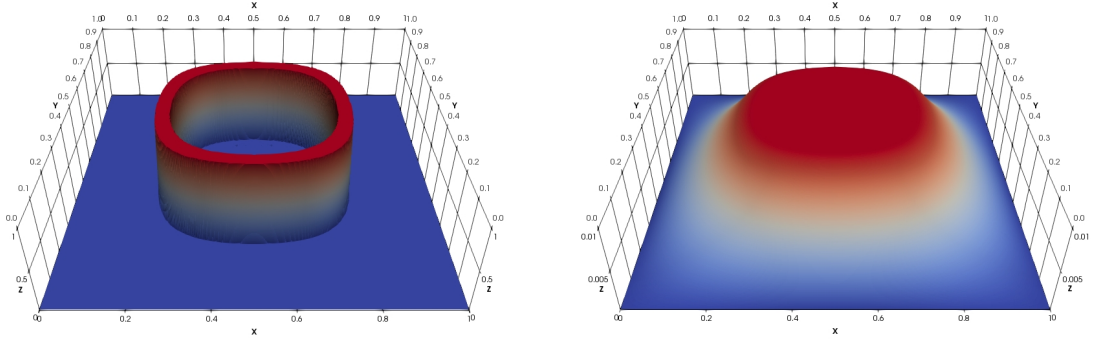
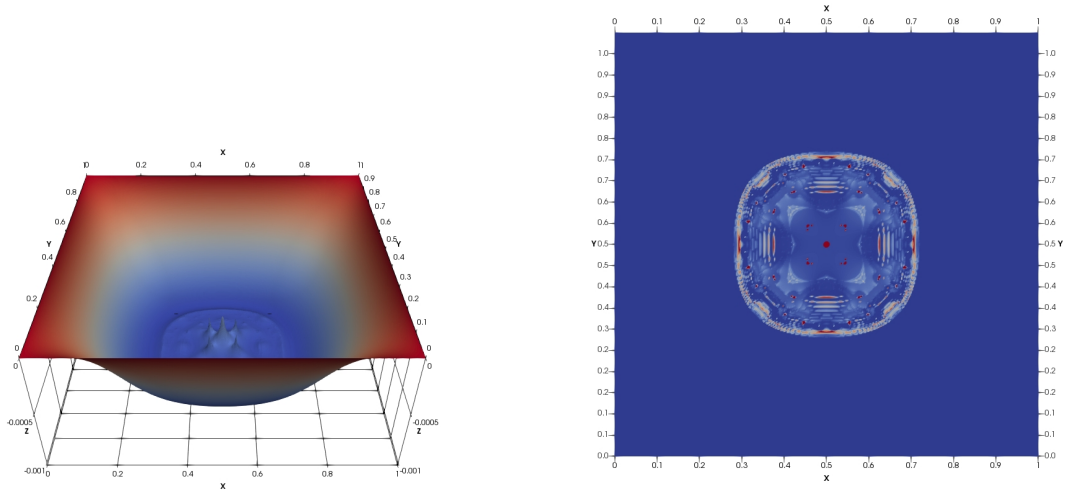Figure 5: Computed results for example 3. From left to right: Control u, state y.

Figure 6: Computed results for example 3. From left to right: Adjoint state p and multiplier μ. The range of μ is given by $\mu(x) \in [0, 40]$

References

[1] M. Bergounioux. Augmented Lagrangian method for distributed optimal control problems with state constraints. *J. Optim. Theory Appl.*, 78(3):493–521, 1993.

[2] M. Bergounioux. On boundary state constrained control problems. *Numer. Funct. Anal. Optim.*, 14(5-6):515–543, 1993.

[3] M. Bergounioux, K. Ito, and K. Kunisch. Primal-dual strategy for constrained optimal control problems. *SIAM J. Control Optim.*, 37(4):1176–1194, 1999.

[4] M. Bergounioux and K. Kunisch. Primal-dual strategy for state-constrained optimal control problems. *Comput. Optim. Appl.*, 22(2):193–224, 2002.

[5] J. F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems*. Springer Series in Operations Research. Springer-Verlag, New York, 2000.

[6] E. Casas. Second order analysis for bang-bang control problems of PDEs. *SIAM J. Control Optim.*, 50(4):2355–2372, 2012.

[7] E. Casas, J. C. de los Reyes, and F. Tröltzsch. Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints. *SIAM J. Optim.*, 19(2):616–643, 2008.

[8] E. Casas, R. Herzog, and G. Wachsmuth. Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional. *SIAM J. Optim.*, 22(3):795–820, 2012.

[9] E. Casas. Control of an elliptic problem with pointwise state constraints. *SIAM J. Control Optim.*, 24(6):1309–1318, 1986.

[10] E. Casas and F. Tröltzsch. Second-order and stability analysis for state-constrained elliptic optimal control problems with sparse controls. *SIAM J. Control Optim.*, 52(2):1010–1033, 2014.

[11] F. H. Clarke. A new approach to Lagrange multipliers. *Math. Oper. Res.*, 1(2):165–174, 1976.

[12] J. C. De los Reyes. *Numerical PDE-constrained optimization.* Springer, Heidelberg, 2015.

[13] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.

[14] O. Güler. On the convergence of the proximal point algorithm for convex minimization. *SIAM J. Control Optim.*, 29(2):403–419, 1991.

[15] M. Hanke and C. W. Groetsch. Nonstationary iterated Tikhonov regularization. *J. Optim. Theory Appl.*, 98(1):37–53, 1998.

[16] M. Hinze. A variational discretization concept in control constrained optimization: The linear-quadratic case. *Computational Optimization and Applications*, 30(1):45–61, 2005.

[17] M. Hinze and U. Matthes. A note on variational discretization of elliptic Neumann boundary control. *Control Cybernet.*, 38(3):577–591, 2009.

[18] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*. Springer, New York, 2009.

[19] K. Ito and K. Kunisch. The augmented Lagrangian method for equality and inequality constraints in Hilbert spaces. *Math. Programming*, 46(3, (Ser. A)):341–360, 1990.

[20] K. Ito and K. Kunisch. Semi-smooth Newton methods for state-constrained optimal control problems. *Systems Control Lett.*, 50(3):221–228, 2003.

[21] M. Jaćimović, I. Krnić, and O. Obradović. On the convergence of a class of the regularization methods for ill-posed quadratic minimization problems with constraint. *Publ. Inst. Math. (Beograd) (N.S.)*, 97(111):89–102, 2015.

[22] V. Karl and D. Wachsmuth. An augmented lagrange method for elliptic state constrained optimal control problems. Preprint SPP1962-008, 2017.

[23] A. Logg, K.-A. Mardal, G. N. Wells, et al. *Automated Solution of Differential Equations by the Finite Element Method.* Springer, 2012.

[24] A. Logg and G. N. Wells. Dolfin: Automated finite element computing. *ACM Transactions on Mathematical Software*, 37(2), 2010.

[25] B. Martinet. Brève communication. régularisation d'inéquations variationnelles par approximations successives. *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, 4(R3):154–158, 1970.

[26] F. Pörner. A priori stopping rule for an iterative Bregman method for optimal control problems. *Optimization Methods and Software*, 2016.

[27] F. Pörner and D. Wachsmuth. An iterative Bregman regularization method for optimal control problems with inequality constraints. *Optimization*, 65(12):2195–2215, 2016.

[28] R. T. Rockafellar. Augmented Lagrangians and applications of the proximal point algorithm in convex programming. *Math. Oper. Res.*, 1(2):97–116, 1976.

[29] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Control Optimization*, 14(5):877–898, 1976.

[30] G. Stadler. Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices. *Comput. Optim. Appl.*, 44(2):159–181, 2009.

[31] F. Tröltzsch. *Optimal control of partial differential equations*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.

[32] D. Wachsmuth and G. Wachsmuth. Regularization error estimates and discrepancy principle for optimal control problems with inequality constraints. *Control Cybernet.*, 40(4):1125–1158, 2011.

[33] D. Wachsmuth and G. Wachsmuth. Necessary conditions for convergence rates of regularizations of optimal control problems. In *System modeling and optimization*, volume 391 of *IFIP Adv. Inf. Commun. Technol.*, pages 145–154. Springer, Heidelberg, 2013.

[34] G. Wachsmuth and D. Wachsmuth. Convergence and regularization results for optimal control problems with sparsity functional. *ESAIM Control Optim. Calc. Var.*, 17(3):858–886, 2011.